DOCUMENT RESUME

ED 110 477 TM 004 750

AUTHOR Wexler, Norman
TITLE Concurrent Validity of the National Teacher Examinations.
PUB DATE Mar 75
NOTE 12p.

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS College Students; *Correlation; *Grade Point Average; Higher Education; Seniors; *Teacher Education; Test Results; *Tests; *Test Validity
IDENTIFIERS *National Teacher Examinations, NTE

ABSTRACT

To investigate the concurrent validity of the National Teacher Examinations, test scores of over 31,000 candidates were correlated with self-reported Grade Point Averages (GPA's). The overall correlation between the Weighted Common Examination Total (WCET) and GPA was .37. Validity indices for the Area Examinations ranged from .08 to .50 with a median of .33. Using 18 selected institutions, the correlation of their mean WCET scores within five GPA levels and GPA level was .70. The latter result suggests overall correlations behave more like lower bound estimates. The WCET and most Area Examinations were concluded to have at least moderate concurrent validity. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED110477

CONCURRENT VALIDITY OF THE NATIONAL TEACHER EXAMINATIONS

U S DEPARTMENT OF HEALTH EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM 004 750

March, 1975

Norman Wexler

One of the major problems in an empirical study of test validity is that of obtaining accurate criterion data. In many instances one must rely on slow and costly questionnaires and be contented with whatever responses received. The National Teacher Examination (NTE) program at Educational Testing Service (ETS) has attempted to deal effectively with this problem in efforts to conduct empirical studies relating to the quality of its tests. Starting with the November, 1973 administration of the NTE, research information has been collected directly from candidates by means of a series of questions printed on the registration form. The questions are both biographical and educational and candidates are assured that their responses will not affect test scores. Subsequently, the responses can be matched with candidate scores (for research purposes only) and analyzed so as to study certain technical aspects of the tests or program.

The present study developed out of an initial investigation of the correlations between the candidate background data and performance on the NTE. The results revealed a positive correlation between candidate self-reported grade point average and NTE performance. This paper represents a detailed description of that finding. Although this study yields substantial information pertaining to the concurrent validity of the NTE, it was not initially designed as a definitive study of the NTE's concurrent validity.

PURPOSE:

The purpose of the National Teacher Examinations is to objectively assess the academic preparation of college seniors. The Common Examinations provide an appraisal of a prospective teacher's basic professional preparation and of representative aspects of general educational studies. The Professional Education Test (110 items) measures achievement in three dimensions of professional education: psychological foundations; societal foundations; and teaching principles and practices. The three General Education Tests are: Written English Expression (45 items); Social Studies, Literature, and the Fine Arts (65 items); Science and Mathematics (50 items). The Weighted Common Examinations Total (WCET) is a sum of scores on the four tests weighted, respectively, 4, 1, 2.5 and 2.5. The twenty-eight Area Examinations aid in evaluating the candidates' preparation to teach or practice in their chosen fields.

The content validity of the NTE is discussed in detail in National Teacher Examinations Technical Handbook (ETS, 1973) which contains references to related empirical studies. The purpose of this study, however, is to estimate the concurrent validity of the NTE using self-reported cumulative undergraduate grade point average (GPA) levels (of 5 ranges) as the criterion measures and Pearson-product-moment correlation coefficients between the WCET and GPA level, and between an Area Examination score and GPA level as indices of concurrent validity. The study was conducted under the assumption that grade point averages reflect academic success, both in the professional and general components of teacher education curricula and in the various areas of teaching specializations.

A number of subscores are normally computed for each candidate taking the NTE; however, this study focused on just those scores (WCET and Area Examination score) upon which designated receivers of the scores typically make decisions and interpretations. (The NTE subscores are substantially correlated with the WCET; see the NTE Technical Manual).

SAMPLE:

The sample for the present study consisted of all candidates taking the NTE in November 1973 and January 1974 (over 30,000 candidates). The number of subjects ranging from 25 to 10,036 varied from one result to another depending on several factors, such as the particular Area Examination taken and the availability of self-reported candidate information.

It should be noted that the test-taking population of the NTE is not geographically representative of all persons entering the teaching profession. Approximately 75% of the NTE candidates are from the South Central, South and Middle Atlantic, and the New England States. There was, however, no reason to suspect *a priori* that there would be systematic differences in the relationship of NTE scores to GPA level between beginning teachers from the eastern part of the United States and other beginning teachers. No such investigation was subsequently made nor deemed necessary.

PROCEDURES:

Correlations were obtained between self-reported GPA and both the WCET and the Area Examinations. The correlations were separately computed by area and also for all candidates. In addition, for each area, the correlation

between the WCET and Area Examination score was obtained using all candidates with both scores available.

Ordinarily, one would expect that a grade point average is at least in part a function of the institution attended. Thus, as the basis for another correlation in the study, scores from a sample of educational institutions were drawn. It was judged that the institutions should be reasonably large in order to promote stability of the data and integrated to reduce the chance of racially biased data (approximately 20% of the NTE candidates are Black). Hence, the candidate score files for the pooled administrations were searched in such a manner so as to select all scores from any institution for which there were at least 20 Black and 20 White candidates who designated the institution as their undergraduate school. Eighteen such institutions were detected. The WCET scores for all students who were members of the eighteen selected schools were processed so as to produce the mean WCET for each level of GPA within each of the selected institutions. These derived means were then used as data to compute a correlation with GPA. The effect of this procedure was to observe the relationship between the self-reported GPA level and the WCET--allowing only institutional variation to operate within GPA level.

RESULTS:

Table I reports means, standard deviations, and concurrent validity indices for 24 of the 28 Area Examinations administered nationally. Data for the tests in German, Introduction to the Teaching of Reading, Texas Government, and Audiology are not reported since there were fewer than twenty candidates available for those areas. The validity indices ranged from .08 to .50 with a median index of about .33 or .34. It is not clear why some tests have a relatively low index or, conversely, a relatively high index. Areas such as Guidance Counselor, Educational Administration, and Reading Specialist could be expected to have lower indices because they are reflections of graduate programs in which, typically, candidates have high undergraduate GPA's which do not vary nearly as much as those of candidates who do not enter graduate programs; hence, the validity indices tend to be lower. This explanation does not, however, account for low indices for the tests in Men's Physical Education, Chemistry, Physics, and General Science, or Education in an Urban Setting. It is similarly puzzling why some tests produce a high index.

.....

Insert Table I about here

.....

Correlations between GPA level and WCET by area are about the same magnitude as the validity indices for the Area Examinations. These are shown separately in Table I and have a range of .16 to .47 with a median correlation of about .36. The correlation between GPA level and WCET using all students regardless of their teaching area was computed to be .367 and is consistent with the median of the correlations reported separately by area. This relationship is displayed in Figure 1 which shows selected percentiles of the distribution of WCET by GPA level.

It is important to keep in mind that the relationship depicted in Figure 1 is a result which ignores the variation in GPA from one institution to another; for example, a 3.0 GPA in one school does not necessarily indicate the same level of ability or achievement as a 3.0 in another school. Despite this possible confounding, there are differences ranging from 23 to 47 scale score points between the means at successive self-reported GPA levels.

.....

Insert Figure 1 about here

.....

Correlations between the WCET and the Area Examinations are reported in Table II. These correlations are fairly substantial, ranging from .40 to .90 with a median coefficient of about .84. These results suggest a considerable amount of overlap between the two scores. It is reasonable, however, to expect that students who do well in their specialties will do well generally and conversely. Nevertheless, an inspection of Table II shows that the intercorrelations are consistently lower than the Area Examination reliabilities. This result implies that some specific variance remains in one or both scores that may offer additional information unrelated to grades.

.....

Insert Table II about here

.....

TABLE I

MEANS, STANDARD DEVIATIONS, AND CONCURRENT VALIDITIES OF THE NTE BY TEACHING OR SUBJECT FIELD

Teaching or Professional Area	WCET Number of Candidates	WCET Mean	WCET S.D.	Correlation GPA & WCET	AREA Number of Candidates	AREA Mean	AREA S.D.	Correlation Between GPA & Area Exams.
Art Education	628	582	80	.31	727	587	89	.34
Biology and General Science	612	601	98	.35	694	635	85	.31
Business Education	948	518	90	.32	1,050	582	86	.29
Chemistry, Physics and General Sci.	168	659	87	.38	191	627	86	.22
Early Childhood Education	3,477	564	97	.38	4,333	612	96	.35
Education in the Elementary Sch.	8,334	571	97	.34	10,036	603	91	.32
Education of the Mentally Retarded	1,228	583	92	.39	1,[illegible]21	605	93	.38
Education in an Urban Setting	25	552	98	.16	30	539	73	.22
Educational Admin. & Supervision	136	598	92	.20	439	613	95	.19
English Language and Literature	1,796	619	91	.38	2,071	588	88	.39
French	168	614	85	.36	216	597	111	.50
Guidance Counselor	64	619	87	.37	242	620	91	.08
Home Economics Education	755	562	81	.41	811	613	84	.38
Industrial Arts Education	244	532	100	.24	302	593	80	.29
Mathematics	992	618	96	.35	1,126	604	91	.34
Media Specialist--Library & A/V	160	600	101	.36	230	603	85	.38
Men's Physical Education	1,080	516	83	.24	1,226	587	83	.22
Music Education	751	576	93	.39	829	594	92	.43
Reading Specialist--Elem. Sch.	62	619	104	.29	209	606	86	.20
Social Studies	2,157	584	105	.31	2,406	584	92	.29
Spanish	222	580	95	.42	340	659	93	.33
Speech--Communication & Theatre	201	577	81	.40	217	568	105	.41
Speech Pathology	233	583	88	.47	988	619	99	.36
Women's Physical Education	854	541	89	.40	963	613	97	.35

TABLE II

RELATIONSHIPS BETWEEN THE WEIGHTED COMMON EXAMINATION TOTAL AND AREA EXAMINATIONS

Teaching or Professional Area	Correlation Between WCET and Area Examination Score	Reliability * of Area Examination
Art Education	.78 (667)**	.92
Biology and General Science	.85 (633)	.95
Business Education	.86 (981)	.91
Chemistry, Physics and General Science	.79 (179)	.94
Early Childhood Education	.89 (3,670)	.90
Education in the Elementary School	.88 (8,812)	.92
Education of the Mentally Retarded	.85 (1,266)	.89
Education in an Urban Setting	.85 (25)	.92
Educational Admin. & Supervision	.79 (147)	.89
English Language and Literature	.90 (1,876)	.94
French	.65 (174)	.94
Guidance Counselor	.81 (77)	.92
Home Economics Education	.86 (774)	.90
Industrial Arts Education	.87 (260)	.92
Mathematics	.71 (1,029)	.93
Media Specialist--Library & A/V	.84 (171)	.94
Men's Physical Education	.83 (1,130)	.87
Music Education	.83 (778)	.91
Reading Specialist--Elem. School	.84 (69)	.90
Social Studies	.89 (2,286)	.95
Spanish	.40 (233)	.94
Speech--Communication & Theatre	.79 (205)	.82
Speech Pathology	.72 (240)	.94
Women's Physical Education	.88 (885)	.91

* From formal test analyses published at ETS.

** Number of candidates in parentheses.

Results of the separate correlation using eighteen selected schools are shown in Figure 2. The points within each GPA level represent the mean WCET at those schools for all candidates whose self-reported GPA's were at that level. The mean of those points is reported above each level designation along with the total number of candidates at the GPA level in all eighteen schools. A difference can be observed between the means at adjacent levels; moreover, the differences are quite marked for the high GPA levels. Using all seventy mean WCET points in the figure and corresponding GPA level as paired data, the correlation between them was found to be .70. This coefficient is much higher than that reported in Figure 1 because the variation among students is removed.

The appreciable variation among the institutional WCET means at fixed GPA levels suggests that if sufficient data were available within individual institutions, the concurrent validities for the WCET would be higher than those shown in Table I. This paper takes the position that the concurrent validities reported in Table I behave more like lower bound estimates of concurrent validity. In other words, typical estimates of concurrent validity would more than likely be higher than those reported in Table I if the data were gathered from within an individual school.

.....

Insert Figure 2 about here

.....

SUMMARY AND CONCLUSIONS:

Using more than 30,000 candidates the WCET of the National Teacher Examination Program had an overall correlation of .37 with self-reported GPA level. Validity indices for the Area Examinations ranged from .08 to .50 with a median index of .33 using all candidates within each area during the November 1973 and January 1974 administrations.

In a separate correlation using institutional WCET means and GPA levels, the index of relationship was found to be .70.

An interpretation was made that the overall indices confounding institutional differences were more like estimates of the lower bound of concurrent validity.

The results discussed here give a reasonable indication that the WCET score and most Area Examinations scores have moderate concurrent validity using a self-reported GPA level as the criterion measure.

Although the GPA does not actually reflect the substance of one's curricular exposure, it is a readily available index among the credentials of a prospective teacher that is usually evaluated by a potential employer. The variation of grading standards among institutions as depicted in Figure 2, however, implies the need for a standard objective instrument from which one can infer and/or corroborate both the content and standing of a candidate's academic preparation.

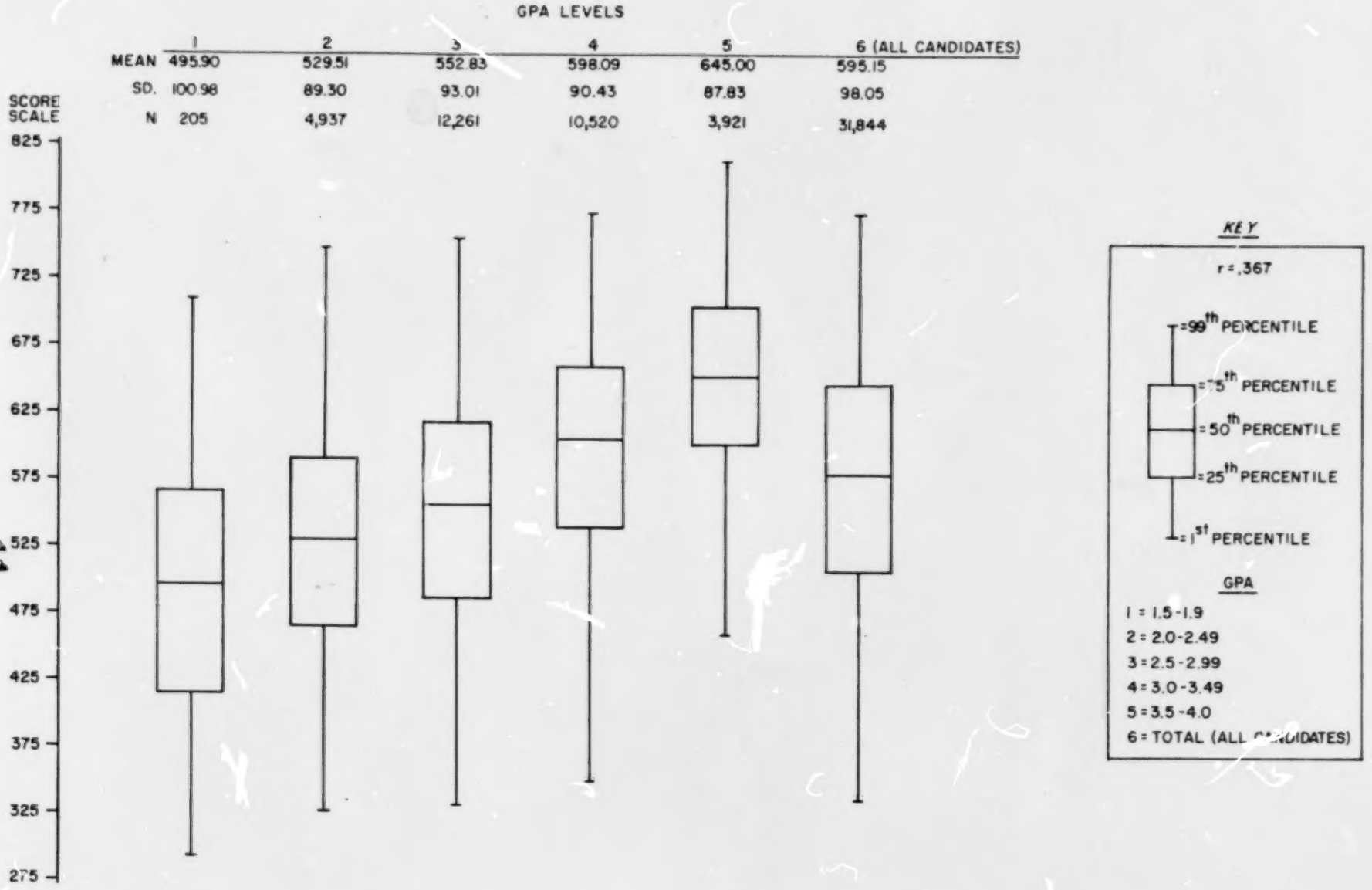

FIGURE I: SELECTED PERCENTILES OF THE WEIGHTED COMMON EXAMINATION TOTAL BY SELF-REPORTED GPA LEVELS

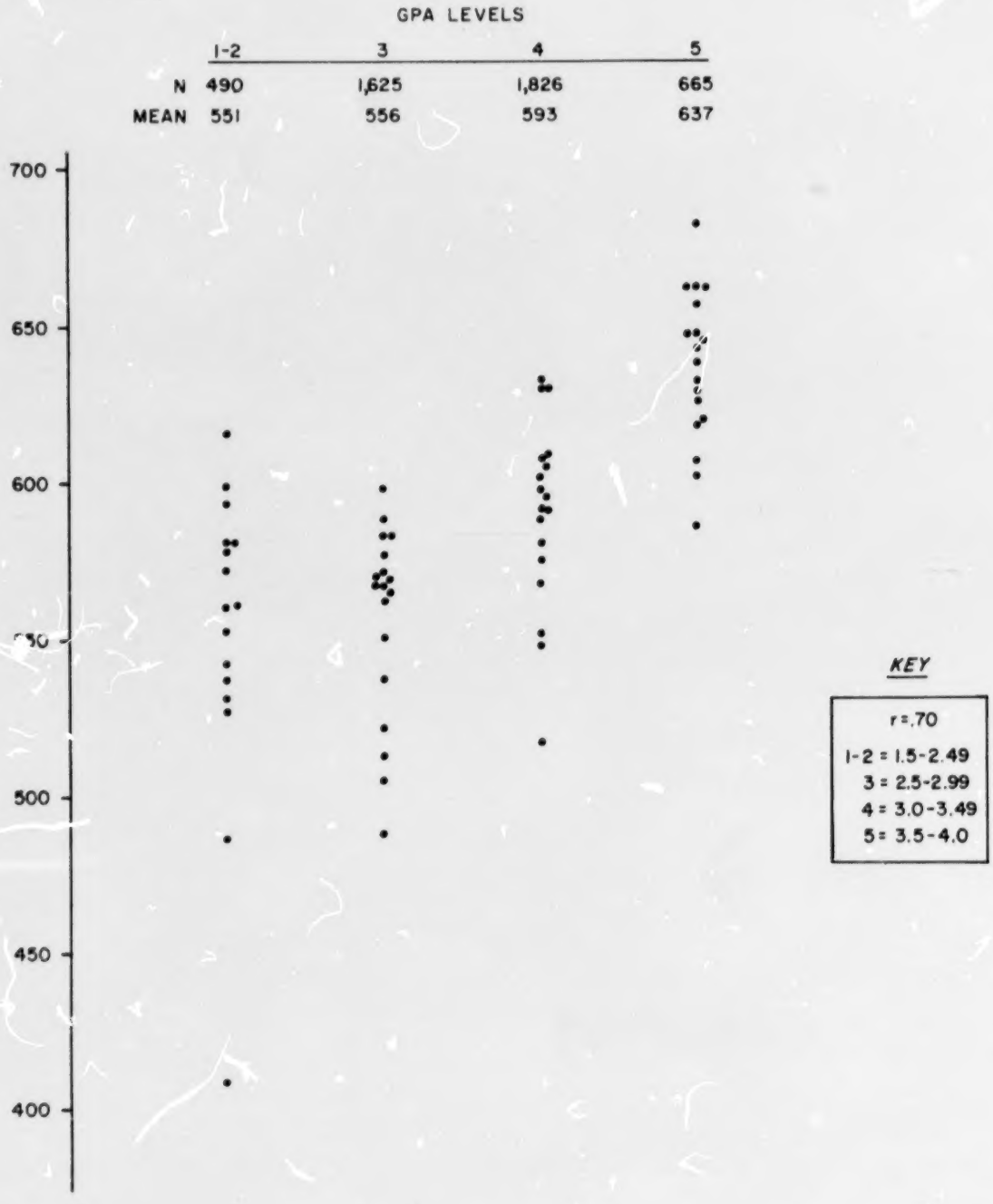

FIGURE 2: WEIGHTED COMMON EXAMINATION TOTAL UNWEIGHTED MEAN SCORES OF EIGHTEEN SELECTED SCHOOLS BY SELF-REPORTED GPA LEVELS*

*SCORES WERE POOLED WITH ADJOINING GPA LEVEL IN THE LESS EXTREME DIRECTION IF LESS THAN TWO CANDIDATES WERE IN A PARTICULAR GPA LEVEL